

AC P09799; DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VICTILIN GC72-A PRECURSOR (ALPHA-GLOBULIN A).
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicots; core eudicots; Rosidae; eurosids II;
 OC Malvales; Malvaceae; Gossypium.
 [1]
 RP SEQUENCE FROM N A.
 RA Chilan C.A., Borroto K., Kamalay J.A., Dure L. III;
 RT Developmental biochemistry of cottonseed embryogenesis and alpha
 globulin (victilin) genes of cottonseed.;
 RL Plant Mol. Biol. 9:53-545(1987).
 CC -!- FUNCTION: SEED STORAGE PROTEIN.
 CC -!- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN
 CC BODIES.
 CC -!- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASPOLIN, VICTILIN,
 CC CONVICTILIN, CONGLYCININ, ETC.).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U20432; AAA62212.1; -.
 DR HSSP; P13569; INPD.
 DR DICTYDB; DDO2059; TAGB.
 DR INTERPRO; IPR002039; -.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR INTERPRO; IPR001617; -.
 DR PFAM; PF00664; ABC_membrane; 1.
 DR PFAM; PF00005; ABC_tran; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASPI; FALSE_NBG.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR KW Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;
 KW SIGNAL 1 23
 FT CHAIN 24 605 VICTILIN GC72-A.
 FT SIGNAL 1 23
 FT CHAIN 24 605 AA; 71049 MW; C9DB9371C976953B CRC64;
 SQ -----
 Query Match 63.1%; Score 142; DB 1; Length 605;
 Best Local Similarity 53.8%; Pred. No. 1.1e-07; Matches 21; Conservative 12; Mismatches 6; Indels 0; Gaps 0;
 OY 2 EDPPQRYYEEQECQEQCQEERRPQQCQCRKRFEQQQ 40
 OY ||||| ||:::|| :|| | ::|| | :|| :|||
 Db 78 EDPPQRYYEEQECQEQCQEERRPQQCQCRKRFEQQQ 116
 RESULT 3
 TAGB_DICDI ID TAGB_DICDI STANDARD; PRT; 1905 AA.
 AC P54683; DT 01-OCT-1995 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE PRESTALK SPECIFIC PROTEIN PRECURSOR (EC 3.4.21.-).
 GN TAGB.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Dictyosteliida; Dictyostelium.
 RN [1]
 RP SEQUENCE FROM N A.
 RC STRAIN_AXI;
 MEDLINE=95262903; PubMed=7744252;
 RA Shaulsky G., Kuspa A., Loomis W.F.;
 RT "A multidrug resistance transporter/serine protease gene is required
 for prestalk specialization in Dictyostelium.";
 Genes Dev. 9:111-112(1995).
 RL -!- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGB MAY MEDIATE
 CC INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO PEPTIDASE FAMILY
 CC S8; ALSO KNOWN AS THE SUBTILASE FAMILY.
 CC -----
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING
 CC TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
 CC -!- SIMILARITY: STRONG, TO TAGB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U20432; AAA62212.1; -.
 DR HSSP; P13569; INPD.
 DR DICTYDB; DDO2059; TAGB.
 DR INTERPRO; IPR002039; -.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR INTERPRO; IPR001617; -.
 DR PFAM; PF00664; ABC_membrane; 1.
 DR PFAM; PF00005; ABC_tran; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASPI; FALSE_NBG.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR KW Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;
 KW SIGNAL 1 23
 FT CHAIN 24 605 AA; 71049 MW; C9DB9371C976953B CRC64;
 SQ -----
 Query Match 38.2%; Score 86; DB 1; Length 1905;
 Best Local Similarity 42.5%; Pred. No. 0.12; Matches 17; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

Db	1809 PPPPQQEQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 1848	Query Match	36.98;	Score	83;	DB	1;	Length	429;			
RESULT	4	Best Local Matches	46.28;	Pred.	No.	0	066;	Indels	0;	Gaps	0;	
ID	AP04_MACFA	STANDARD:	PRT:	429 AA.	QY	2	EDPQRVVECCQQCQRCQFEROQPCQCRCLRFEQHQQQ 40	Db	378 EQQEQQQEQEQQQQEQQQQEQREQQEQQQEQ 416			
AC	P3621;				AC	01-FEB-1994 (Rel. 28, Created)						
DT	01-FEB-1994 (Rel. 28, Last sequence update)				DT	15-DEC-1998 (Rel. 37, Last annotation update)						
DE	APOLIPOPROTEIN A-IV PRECURSOR (APO-A-IV).				DE	APOA4.						
GN	Macaca fascicularis (Crab eating macaque) (<i>Cynomolgus monkey</i>).				GN	Macaca fascicularis (Nucleotide sequences of the Macaca fascicularis apolipoprotein C-III and A-IV genes);						
RL	Medline:9312330; PubMed=8448212;				RL	Blochim. Biophys. Acta 1172:335-339(1993).						
CC	Osada J., Pocovi M., Nicolosi R.J., Schaefer E.J., Ordovas J.M.;				CC	-I- FUNCTION: MAY HAVE A ROLE IN CHYLOMIGRONS AND VLDL SECRETION AND CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR COMPONENT OF HDL AND CHYLOMIGRONS.						
RT	Mammalia; Etheria; Primates; Catarrhini; Cercopitheciniae; Macaca.				RT	-I- SUBCELLULAR LOCATION: EXTRACELLULAR.						
RN	[1]				RN	-I- TISSUE SPECIFICITY: SECRETED IN PLASMA.						
RP	SEQUENCE FROM N.A.				RP	-I- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-HELICAL, AND MANY OF THESE HELICES ARE AMPHITRIC. THEY MAY THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.						
RC	-I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.				RC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).						
CC					CC	EMBL: X68361; CRA48421.1; -.						
CC	PIR: S29565; S29565.				CC	DR: HSSP; P02649; INFO: INTERPRO; IPR000074; -.						
CC	INTERPRO; IPR000074; -.				CC	DR: PF01442; Apolipoprotein: 1. PLAMA; Lipid transport; HDL; Chylomicron; Repeat; Signal.						
KW	BY SIMILARITY.				KW	FT: SIGNAL CHAIN 1 20						
FT	APOLIPOPROTEIN A-IV.				FT	FT: DOMAIN 31 330 X 22 AA APPROXIMATE TANDEM REPEATS.						
FT	REPEAT 33 54 1.				FT	FT: REPEAT 60 81 2.						
FT	REPEAT 82 103 3.				FT	FT: REPEAT 115 136 4.						
FT	REPEAT 137 158 5.				FT	FT: REPEAT 159 180 6.						
FT	REPEAT 181 202 7.				FT	FT: REPEAT 203 224 8.						
FT	REPEAT 225 246 9.				FT	REPEAT 247 268 10.						
FT	REPEAT 269 286 11.				FT	REPEAT 287 308 12.						
FT	REPEAT 309 330 13.				FT	REPEAT 372 420 GLU/GLN-RICH.						
SQ	SEQUENCE 429 AA; 49876 MW; 3D458F551D0DB60C CRG64;				SQ	SEQUENCE 429 AA; 49876 MW; 3D458F551D0DB60C CRG64;						

Query Match	36.98;	Score	83;	DB	1;	Length	429;
Best Local Matches	46.28;	Pred.	No.	0	066;	Indels	0;
Matches	18;	Conservative	10;	Mismatches	11;	Indels	0;
QY	2 EDPQRVVECCQQCQRCQFEROQPCQCRCLRFEQHQQQ 40	PRO_DROME	STANDARD;	PRT;	1403 AA.		
DB	378 EQQEQQQEQEQQQQEQQQQEQREQQEQQQEQ 416	PRO_DROME	STANDARD;	PRT;	1403 AA.		
AC	P29617;	AC	01-FEB-1993 (Rel. 25, Created)	AC	01-FEB-1995 (Rel. 33, Last sequence update)		
DT	01-FEB-1993 (Rel. 25, Last sequence update)	DT	30-MAY-2000 (Rel. 39, Last annotation update)	DT	30-JUL-2000 (Rel. 40, Last sequence update)		
RA	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	RA	"prospero is expressed in neuronal precursors and encodes a nuclear protein that is involved in the control of axonal outgrowth in Drosophila;"	RA	"prospero gene encodes a divergent homeodomain protein that controls neuronal identity in Drosophila.";		
OC		OC		OC			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RC		RC		RC			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			
CC		CC		CC			
RT		RT		RT			
RN		RN		RN			
RP		RP		RP			
RA		RA		RA			

RA Green R.;
 RT "The involucrin genes of the mouse and the rat: study of their shared
 RT repeats.";
 RL Mol. Biol. Evol. 10:1136-1149(1993).
 CC !- FUNCTION: INVOLUCRIN IS A KERATINOCTYE PROTEIN THAT FIRST APPEARS
 IN THE CELL CYTOSOL, BUT ULTIMATELY BECOMES CROSS-LINKED TO
 MEMBRANE PROTEINS BY TRANSGLUTAMINASE. ALL THAT RESULTS IN THE
 FORMATION OF AN INSOLUBLE ENVELOPE BENEATH THE PLASMA MEMBRANE.
 CC !- TISSUE SPECIFICITY: PRESENT IN KERATINOCTYES OF EPIDERMIS AND
 OTHER STRATIFIED SQUAMOUS EPITHELIUM.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: L2819; AAA3930.1; -.
 DR MGD; MGI:96626; IVL;
 DR INTERPRO; IPR002360; -.
 DR PROSITE; PS00795; INVOLUCRIN; 1.
 KW KERATINOCTYE; Repeat;
 SQ SEQUENCE: 467 AA; 54.91% MW: 603 EIE51B435737D CRC64;

RESULT 12

Query Match	Score	DB	Length
Best Local Similarity	34.0%	1	467
Matches	16		
Conservative	33.3%		
Mismatches	9		
Indels	9		
Gaps	1		

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame (ORF) similar to a RT translationally controlled tumour protein, one ORF containing motifs RT also found in plant storage proteins and 13 ORFs with weak or no RT homology to known proteins.";			
RL Yeast 10:563-588(1994).			

YKF4 YEAST

ID	STANDARD;	PRT;	AA.
AC P35732;			738 AA.
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-FEB-1996 (Rel. 33, Last sequence update)			
DE HYPOPHETIC PROTEIN 84.0 KA PROTEIN IN NUP120-CSE4 INTERGENIC REGION.			
GN YKL054C OR YKL308.			
OS Saccharomyces cerevisiae (Baker's yeast).			
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
RN SEQUENCE FROM N.A.			
RP STRAIN=S288C;			
RC MEDLINE=93378723; PubMed=8091862;			
RA Rasmussen S.W.;			
RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the RT FBA1 and YKA2 genes, an open reading frame			

CC	FT	DOMAIN DOMAIN DOMAIN DOMAIN	489 714 892 755	588 982 798 564	ASP-GLD-RICH (ACIDIC). PRO-RICH. ARG/LYS-RICH (BASIC).
CC	FT	CONFFLICT	564	564	E -> D (IN REF. 1).
CC	SEQUENCE	905 AA;	102557 MW;	A287D4A648DD1A35	CRC64;
CC	SQ				
Query Match			33.8%	Score 75; DB 1; Length 905;	
Best Local Similarity			35.9%	Pred. No. 0.63; Gaps 0;	
Matches			14; Conservative 16; Mismatches 9; Indels 0;		
Oy			2 EDPORVEECQCRCQEEFEROQPOCQCRCRKFQEQQ 40		
Db			: I::::: II : I:::II II : :: : I:II 263		
RESULT 14					
CLOC_DROME			STANDARD; PRT; 1023 AA.		
ID			CLOCK DRONE		
AC			061735; 076342; 07137;		
DT			15-JUL-1999 (Rel. 38, Last sequence update)		
DE			CIRCADIAN LOCOMOTOR OUTPUT CYCLES		
GN			KAPUT PROTEIN (DCLOCK) (DPAS1).		
OS			DROSOPHILA MELANOGASTER (fruit fly).		
OC			Bukarvilia; Metrazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC			Pterigota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
RN			Ephydriidea; Drosophilidae; Drosophila.		
RP			[1]		
SEQUENCE FROM N_A.					
RC			TISSUE=HEAD;		
RX			MEDLINE=98279147; PubMed=9616122;		
RA			Darlington T.K., Wager-Smith K., Ceriani M.F., Staknis D., Gekakis N.,		
RA			Steeves T.D.L., Weitz C.J., Takahashi J.S., Kay S.A.;		
RT			"Closing the circadian loop: CLOCK-induced transcription of its own		
RT			inhibitors per and tim."		
RL			Science 280:1599-1603(1998).		
RN			[2]		
RP			SEQUENCE FROM N_A., AND MUTAGENESIS.		
RC			TISSUE=HEAD;		
RC			MEDLINE=98292177; PubMed=9630223;		
RX			Allada R., White N.E., So W.V., Hall J.C., Rosbash M.;		
RA			"A mutant drosophila homolog of mammalian Clock disrupts circadian		
RT			rhythms and transcription of period and timeless.";		
RL			Cell 93:791-804(1998).		
RN			[3]		
RP			SEQUENCE FROM N_A.		
RC			STRAIN=CANTON-S;		
RX			MEDLINE=98414630; PubMed=9742131;		
RA			Bae K., Lee C., Sidore D., Chuang K.-Y., Edery I.;		
RT			"Circadian regulation of a Drosophila homolog of the mammalian clock		
RT			gene: PER and TIM function as positive regulators.";		
RL			Mol. Cell. Biol. 18:6142-6151(1998).		
CC			-!- FUNCTION: CIRCADIAN REGULATOR THAT ACTS AS A TRANSCRIPTION FACTOR		
CC			AND GENERATES A RHYTHMIC OUTPUT WITH A PERIOD OF ABOUT 24 HOURS.		
CC			OSCILLATES IN ANTI-PHASE TO THE CYCLING OBSERVED FOR PERIOD (PER)		
CC			AND TIMELESS (TIM). ACCORDING TO REF.3, REACHES PEAK ABUNDANCE		
CC			WITHIN SEVERAL HOURS OF THE DARK-LIGHT TRANSITION AT ZT0		
CC			(ZEITGEBER 0), WHEREAS REF.1 DESCRIBES BIMODAL OSCILLATING		
CC			EXPRESSION WITH MAXIMUM AT ZT5 AND ZT23. CLOCK-CYCLE HETEROODIMERS		
CC			ACTIVATE CYCLING TRANSCRIPTION OF PER AND TIM BY BINDING TO THE E-		
CC			BOX (3'-CACGTG-5') PRESENT IN THEIR PROMOTERS. ONCE INDUCED, PERIOD AND TIMELESS BLOCK CLOCK'S ABILITY TO TRANSACTIVATE THEIR		
CC			PROMOTERS.		
CC			-!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER		
CC			BHLH PROTEIN. FORMS A HETEROODIMER WITH CYCLE.		
CC			-!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).		
CC			-!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; THE FULL-LENGTH VARIANT A (SHOWN HERE) AND VARIANT B; ARE PRODUCED BY ALTERNATIVE SPLICING.		
CC			VARIANT B ENCODES TWO CONCEPTUAL PROTEINS, THE FIRST CONSISTS ONLY		
CC			OF THE BHLH DOMAIN, THE OTHER CONSISTS OF THE PAS-1 AND ALL C-		
CC			TERMINAL DOMAINS. VARIANT B IS EXPRESSED WEAKLY AT ALL THE TIME OF		
RESULT 15					
WC1_NEUCR			STANDARD; PRT; 1154 AA.		
ID			WC1_NEUCR		
AC			001371;		
DT			01-NOV-1997 (Rel. 35, Created)		
DT			01-NOV-1997 (Rel. 35, Last sequence update)		
DT			30-MAY-2000 (Rel. 39, Last annotation update)		
QY			5 QRRVEECQCRCQEEFEROQPOCQCRCRKFQEQQ 40		
Db			: I::::: II : I:::II II : :: : I:II 829		
Query Match			33.3%	Score 75; DB 1; Length 1023;	
Best Local Similarity			38.9%	Pred. No. 0.88; Gaps 0;	
Matches			14; Conservative 15; Mismatches 7; Indels 0;		
Oy			5 QRRVEECQCRCQEEFEROQPOCQCRCRKFQEQQ 40		
Db			: I::::: II : I:::II II : :: : I:II 829		

DE WHITE COLLAR 1 PROTEIN (WC1).
 GN WC-1.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=74-OP23-1A;
 RX MEDLINE#95203083; PubMed#8612589;
 RA Ballario P., Vittorioso P., Magrelli A., Talora C., Cabibbo A.,
 Macino G.;
 RT "White collar-1, a central regulator of blue light responses in
 Neurospora, is a zinc finger protein.";
 RL EMBJ J. 15:1650-1657(1996).
 CC -I- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR INVOLVED IN LIGHT
 REGULATION. BINDS AND AFFECTS BLUE LIGHT REGULATION OF THE AL-3
 GENE. WC1 AND WC2 PROTEINS INTERACT VIA HOMOLOGOUS PAS DOMAINS,
 CC BIND TO PROMOTERS OF LIGHT REGULATED GENES SUCH AS FRQ, AND
 ACTIVATE TRANSCRIPTION.
 CC -I- SUBUNIT: HETEROODIMER OF WC1 AND WC2 (POTENTIAL).
 CC -I- SUBCELLULAR LOCATION: NUCLEAR.
 CC -I- INDUCTION: BY BLUE LIGHT.
 CC -I- DOMAIN: THE GLUTAMINE-RICH DOMAIN MIGHT FUNCTION IN ACTIVATING
 GENE EXPRESSION.
 CC -I- SIMILARITY: BELONGS TO THE GATA-TYPE TRANSCRIPTION FACTOR FAMILY.
 CC -I- SIMILARITY: CONTAINS A PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X94300; CAA63964.1; - .
 DR HSSP; P17678; IGRU.
 DR INTERPRO; IPR00014; - .
 DR INTERPRO; IPR000679; - .
 DR INTERPRO; IPR00161; - .
 PFAM; PF00320; GATA; 1.
 DR PFAM; PF00785; PAC; 1.
 DR PFAM; PF00989; PAS; 3.
 DR PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
 DR PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
 KW Transcription regulation; Activator; DNA-binding; Zinc-finger;
 FT Nuclear protein; Repeat.
 FT DOMAIN 16 61 GIN-RICH.
 FT REPEAT 381 448 PAS-1.
 FT DOMAIN 469 508 PAC MOTIF.
 FT REPEAT 576 642 PAS-2.
 FT DOMAIN 650 691 PAC MOTIF.
 FT REPEAT 695 761 PAS-3.
 FT ZN FING 935 960 GATA-TYPE.
 FT DOMAIN 21 57 POLY-GIN.
 FT DOMAIN 329 333 POLY-PRO.
 SQ SEQUENCE 1154 AA; 125944 MW; CB4CB26047ECCDC5 CRC64;

Query Match 33-38: Score 75; DB 1; Length 1154;
 Best Local Similarity 38.5%; Pred. No. 0.98; Indels 0; Gaps 0;
 Matches 15; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Qy 1 PRDPQRYYECQQCQRCQERQOPQCQCRKRFEGEQQ 39
 Qy ||: | : || : ||: ||:|| | ||: :: :||
 Db 12 PEELOHOMHQHQDQQDQQDQQDQQDQQDQQDQQDQQDQQ 50

